

10-9-67

Notes regarding CRIS-1, Part 4

"INSTRUCTIONS FOR PREPARING REPORTS TO CRIS," CSRS-OD-1258, 7/20/67,
was forwarded with Director Kelly's letter of August 16, 1967.

In 1967 we have received a supply of blank Part 4--or shall we call
them Form 4?? In later years, CSRS will fill in this form down to
Field 85.

WORK UNIT/PROJECT NO.: Example - Calif(Berkeley) MH-2294

REGIONAL PROJECT NO.: Leave blank for all except Regional Research.

84 PERIOD COVERED: Fill in "1967."

CRIS ACCESSION NO.: Leave blank.

TITLE: Give full title, using only the regional title on Regional
research.

PERFORMING ORGANIZATION: Example: Dept of Agricultural Economics
etc.
University of California
Berkeley
Calif 94720

COUNTRY: Leave blank.

KEYWORD BANK: Leave blank.

85 PROGRESS REPORT: Limit progress report to approximately 20
lines which will be 1,600 characters, or 80
characters to one line. Letter or space is a
character.

The machine will automatically cut off anything
beyond 2,400 characters which will include the
OBJECTIVES and APPROACH, Fields 24 and 25, of
CRIS-1, Part 1.

86 CHARACTER COUNT: Leave blank.

87 PUBLICATIONS: Limit to 5.

88 NUMBER OF ADDITIONAL PUBLICATIONS PROCESSED: State any number beyond
the five listed in 87.

89 KEYWORD CHANGES: Leave blank.

90 MAJOR ACHIEVEMENT ...ANTICIPATED...: Check one only in blocks A to G.

TITLES, ABSTRACTS AND KEY WORDS FOR INFORMATION RETRIEVAL FROM CRIS

In cooperation with the research agencies of the USDA and the State agricultural experiment stations, plans for establishing a Current Research Information System are going steadily forward. While a prospective contractor is working out the technical aspects of the system there will be time for our respective research staffs to condense and revise the old abstracts of active projects and prepare new ones for putting them into the "data bank" on tape. Along with the revision of the abstracts or preparation of new ones, a list of "key words" or indexing terms for each project is necessary for retrieval of specific information from the system.

There is room for neither cliches nor effusive expression in an automated system. We all must learn to construct titles and abstracts of barebones simplicity and clarity. We must sacrifice details of interest when necessary to assure inclusion of all the more significant points in limited space. Clarity and specificity are essential. They will help assure that your project will come to the attention of the maximum number of people who should know about it.

TITLES

Try to keep each title within one line of 80 spaces of typewriting. One hundred spaces is permissible. It is neither feasible nor necessary to include in the title all or most of the key words. The report forms will have space for key words, which will afford the real means of retrieving information based on subject matter.

Avoid such expressions as "investigations of," "studies on," "research on," and "evidence on." These words take up space and are unnecessary. Instead of long series of narrow terms, use fewer and broader terms in a title. The title may properly appear "clipped" here, in contrast to the way it should appear upon publication, so long as it is literate and informative.

ABSTRACTS

The total space available for statements of "objective," "plan of work" and "progress" will be the equivalent of 30 lines of 80 typewriter spaces each, or 2400 "characters." This amounts to about 300 words of average length -- but the limitation to 2400 characters must be observed. "Objective" and "plan of work" probably can average from 8 to 12 lines of 80 spaces, but must not exceed 20. The equivalent of no less than 10 lines of 80 spaces must be reserved for progress. More than 300 characters may be used for progress so long as the total for objective, plan and progress is not over 2400.

Under "objective" one can incidentally expand the information for which there might be too little space in the one-line or 100-space title. Three lines is generally enough for a well-considered statement of "objective" of a project that embraces a reasonably limited scope.

For the "plan of work" up to 15 or 16 lines of 80 spaces is available, if needed. This much, however, may too seriously limit space for text on progress. Adhere to brief, tightly constructed expressions that will

require the minimum number of lines actually necessary to present the salient points. Leave as much space as feasible for progress.

PROGRESS

Upon entry of a newly approved project, the space for progress obviously must remain blank until time for entry of the first annual (or semi-annual) progress report.

All projects that have been active for a year require a progress or status report in the initial entry into the data bank.

It is possible to comment on degree of progress, kind of accomplishment, or completed segments or phases of a project, to indicate its status without revealing information not yet ready for disclosure.

The right to defer certain disclosures, however, should not be misused as an excuse for withholding information on degree of progress or status of the work. Lack of objectivity in this report will be unhelpful to both the investigator and to others.

A status report in as little as 300 spaces may be difficult but it is often quite feasible. If you may need more space, hold down the initial "object" and "plan."

KEY WORDS

The effectiveness of either manual or automated systems for storage and retrieval of information depends very largely upon the wise choice of "key words" or indexing terms for titles and abstracts. Nowhere do the niceties of language become more important. There are no neat formulas by which we can choose effective terms. There are, however, a few common-sense principles that will help us to be orderly and consistent in the task.

These key words must have the capacity to guide directly to your particular project any inquirer who should know about it; and to any other projects similar or related to it. Some broad terms as well as narrow ones should be included, however, to facilitate the searching or sorting of projects by larger groups, as though by a coarser screen, when that is desirable. More about this later.

Our language is crowded with words each of which has many meanings. When removed from context such a word may have little or no value as an index term unless it is coupled with one -- or even two other words. Seek key words or index terms with restricted and specific meanings that are widely used with the same meanings in related fields as well as in your own. When such carefully chosen terms are grouped with other suitable terms as a complex, the total key can lead to the desired information with a high degree of success, and to a minimum of unwanted data.

Try for effective one-word terms as "key words" but use two-word terms if it is truly necessary. A two-word term should be commonly accepted and well understood in your field -- and preferably in related fields -- just as a one-word term should be. Avoid double meanings.

Many words of specific meanings represent such broad concepts that they are quite useless for indexing. "Experiments" "investigations" "changes" "differences" "reactions" and "evidence" are some obvious examples. Avoid them as key words; and, if feasible, avoid them in abstracts.

An important aid in choice of key words of maximum usefulness is to think (and write) in good index terms as much as you can, in preparing the brief "title," "objectives" and "plan of work" for your project record. . . Try to choose each word so that collectively they reveal as much specific information as feasible about the problem, materials, organisms, variables and techniques involved.

In calling for information from the system, it is necessary that the programmer use key terms in exactly the same form they were put into the system. It is also necessary to avoid cluttering a system with an unmanageable number of key terms. Therefore, the operators of the system must develop a set of "accepted" terms or "vocabulary" for identification purposes. The National Agricultural Library is developing such a list. CRIS is a collaborator in the project. Not everyone who prepares data on a project for recording on tape will have a copy of this "vocabulary" or "thesaurus" to guide him in his choice of terms. Furthermore, these vocabularies or thesauri, as published at any one time cannot be complete. Terms must be added as researches require.

Those who prepare data for the system can contribute materially to its smooth and effective operation by careful choice of form of key terms. By so doing they will increase the probability of using a term in the exact form that it is already listed in the vocabulary, obviating any editorial change. Or, if extractors and indexers come forth with terms that are not already in the vocabulary it will be helpful if the forms of those terms agree with the forms adopted for the vocabulary. Those terms may then be added to the vocabulary when it is revised; and in the meantime can be used provisionally within CRIS if necessary.

The attached except from the National Agricultural Library's "Manual for Development of the Agricultural/Biological Vocabulary" will be a helpful guide to choice of terms and preferred forms.

WHAT KINDS OF KEY TERMS AND HOW MANY

Each "project" or "work unit" will be coded according to the RPDES and PPBS classifications by members of the staff concerned with administration of research. Those classifications are designed mainly for certain management purposes. They are not selective enough to yield information on narrow, very specific subjects, specific diseases, organisms, or concepts.

For example RPDES A5-B12-C4 refers to the broad subject of parasites of insects affecting vegetables, which might include far too many projects to be helpful in a specific inquiry. Many additional key words are needed, such as names of insects, parasites, and vegetables involved; names of major methods or processes; the major variables or factors, the effects of which are under study.

D-029872

the full-scale development, and finally, the final stage.

is attached. Also attached is the "Appendix C" of GPO document from

an example of a test, detection, and key word which is attached.

The cover of the procedure is determined by the second, depending upon one of the following importance of the "second", be listed first and second. Beyond this, the order of presentation is

The kind of board tested in the preceding cases 1 and 2 might well

more than 15 cents.

descending order of that importance to an interface. In the event that

to will often suffice. He suggests you take some to your idea of the

up to 10 cents, if widely chosen, can prove highly effective. This is

of highly accurate times.

4. Preparation, function, reaction, process, product or construction of

one of more than several of individual interest to scientists.

system of continuity involved, it may be there alternative of the procedure

3. Tools, instruments, equipment, processes, type of approach, or

study.

entity, action, or behavior that is the best object of study of

2. The term "annual" means, among other things, "each year", "every

development, each year, "each year", "each year", "each year",

extra coverage, measurement, coverage, such as: only design,

1. This broad objective of field of interest, such as: only design,

be able to handle certain types of situations.

described at group the best in their own words.

mechanism of an assembly, "physiology" of population, as many individuals of

function of the subassembly of physiology, "whole animal" response,

name) of the same organism, name of major mode under study; some other

targeted against escape. In addition was need such, these are: the name (or

and other examples; A7-B27-C12 reduce bodily to physiology of most that

September 1956, The Second Annual Meeting, "Molecular Biology and Biochemicals".

/ /

- a. "G.," read, "now, today, soon, etc.", etc.
- b. The "Chargaff's form will be read when the name applies to "now today," etc.
- c. The plural form to be pronounced when the name may be read in another

d. Singular and plural.

Here we have three possibilities for reading.

3. Non-form. The non-form is preferred to the plural form, G.

name and vice versa.
to ensure if the name is enunciated under a definite article Diseases. Cross-references may be necessary from disease

antonyms and antonym be read.

fold scope, all pictures.
prokaryote, vegetable, vegetable (including green)
monocotyledonous plants; nucleic acids, extract, extract, etc.

b. Common names will be read first

atom, particle.
pathogenic organisms; bacteria, fungi, microorganisms
proteins and the like.

plasma, amebae, ciliates, paramecia, dinoflagellates.

fungi. Cross-references will be made from common names to
specific names.

a. Semantically names will be read first:

Geological and common names.

biology will be avoided.
diseases, and interlocking vocabularies. Using an auxiliary column,
synthetic organic acids, and the various substances such as oligopeptides,
The applicability of terms will be determined from the definition of

c. Definition of biochemical processes and metabolism.

b. Relative homology, the closer terms first, have been selected previously.

a. Relative homology of one in the literature.

2. Same selection. In general, necessary to begin with the more

- c. The singular will be used for specific properties or conditions, e.g., temperature, viscosity, etc.
- d. Both singular and plural may be required if they represent different concepts, e.g., sulfur and sulfates; chromatic and chromites.
5. Distant Proper. In general, terms consisting of more than one word will be listed in natural word order, rather than inverted, e.g., parthenocarpic plants, not plants, parthenocarpic.
6. Combined Terms. A single concept is often expressed by the combination of individual terms.
- a. A precombined term should be used when the concept which it represents is encountered frequently in the literature.
- Mixed forests
Mineral springs
Plant biochemistry
Social insects
- b. Individual terms are preferred if the combination expresses a relationship reflected by: effect of, effect on, in, an, from, for, and, by.
- Might and Chick embryos, not Might, Effect on chick embryos.
- Selenium and Soils, not Selenium in Soils.
- Pesticides and Insect control, not Pesticides for Insect control.
- Alcohols and Fatty acids, not Alcohols from fatty acids.
- c. Individual terms should be provided for a process and its related device or material.
- Milk and Pasteurization, not Pasteurization of milk.
- Silk and Spinning, not Spinning of silk.
- Beneficiation and Minerals, not Beneficiation of minerals.
7. Term Definition. Terms that have more than one accepted meaning or that must be distinguished from other terms should be accompanied by an explanation. Ordinarily, the cross-reference array will be adequate to place a term in its proper context. Occasionally, other techniques are needed.
- a. Homographs are distinguished by adding a parenthetical explanation after the term, e.g., gold (disease) and gold (impurity); scales (animal) and scales (mechanical); scales (leaves) and scales (weighting instrument). The qualifying explanation is a part of the term.

- b. Other ambiguous terms may have a scope note to explain the proper use of the term. A scope note is enclosed in parentheses but follows the term on a succeeding line and is not a part of the term, for example:

Water cooling
(cooling by water)